Exploring three views on image enhancement for Pixel Privacy

Simon Brugman  
Radboud University, the Netherlands  
simon.brugman@cs.ru.nl

Maciej Wysokiński  
Universidad Complutense, Spain  
maciwyso@ucm.es

Martha Larson  
Radboud University, the Netherlands  
m.larson@cs.ru.nl

ABSTRACT
The aim of the MediaEval 2018 Pixel Privacy task is to increase image appeal while blocking automatic inference of sensitive scene information. We investigate three different views from which we could consider enhancement: the view of the image aesthetics field, the view of automatic large-scale aesthetics inference models, and the view of social media users who reflect on their own photographic practices. Systematic image editing can do better than one-size-fits-all-filters with helping casual social media users find the desired photo look. Machine learning aesthetics assessment falls short when inferring individual preferences. A qualitative user study gives insight into the diversity and complexity of preferences.

1 INTRODUCTION
The MediaEval Pixel Privacy task aims at protecting users from large-scale inference of sensitive information while increasing image appeal. As we develop Pixel Privacy technologies, we want to understand how to apply and assess image enhancement. In this paper, we consider three views on image enhancement.

• The view from the field of image aesthetics: Here we explore what aspects of overall colour harmony we can systematise without full understanding of the content of the image.
• The view of machine learning on automatic inference of aesthetics: We would like to better understand the potential of this technology for aesthetics evaluation of image enhancement in the Pixel Privacy task.
• The view of social media users: We survey a small group of participants who have the habit of consciously reflecting on their own photographic practices. This qualitative user study aims at discovering strong and weak points of our image enhancements.

In the following sections, we discuss each view in turn. Note that in this work we assume an interconnection between enhancement and appeal. Consistently with [11, 14] we consider that improving aesthetics also improves appeal.

2 SYSTEMATIC IMAGE EDITING
We consider the field of image aesthetics in order to discover aspects of photos that can be changed systematically, leading to an aesthetic improvement or an increase of appeal without full knowledge of what is being depicted in the photo. Such aspects would lend themselves well to automation. Our interest in automation is related to the observation that automatic filters are currently in widespread use and assume that transformations must be fast to match the speed of what is currently offered by apps.

The number of amateur photographers is growing as smartphone usage increases [16]. Popular camera mobile apps attract a large amount of activity, such as Instagram (more than 1B monthly active users [15]) and Flickr (the iPhone is the most used camera [7]). These apps allow users to edit images internally, for example, applying filters, supporting extremely fast sharing of edited images.

Currently, the state of the art in mobile apps for cameras is predefined filters, which can change in hue, saturation or lightness or add visual effects like blur or noise. Filtered photos, especially with increased colour temperature, exposure, and contrast, are more likely to be viewed (+21%) and commented on (+45%) than unfilled photos [2]. These filters have the disadvantage of depriving users of editing control. Predefined filters are the same each time the filter is used and may limit the ability of users to achieve the desired photo look. Here, we aim to discover contributions from the field of image aesthetics that would allow us to improve the flexibility of photo filters to increase image aesthetics and add user appeal.

The users of image sharing networks can be divided into people with aesthetic knowledge and casual photographers. The former group tends towards smooth changes, supported by manual editing, the latter usually prefers to achieve more dramatic change [2]. Our goal is to discover dramatic changes consistent with image content, but not requiring full image understanding. Early explorations have directed our attention to colour grading and cropping for image enhancement. Figure 1 shows a colour transformation whose goal is to increase the appeal of the original image. The example was chosen because it is one of the promising cases where the classifier used in the Pixel Privacy task [9] is misdirected by a transformation.

What aspects of overall colour harmony can we systematise without full image understanding? As an initial attempt, we convert the input image to HSV colour space [19], obtaining pixel values expressed in terms of the three-dimensional nature of human colour perception [20]: (1) hue, which refers to pure colour, (2) saturation from white light to pure colour and (3) value, which refers to illumination values. Assigning the hue values to the specific ranges in the RGB colour wheel [13] (primary, secondary and tertiary colours) it is possible to identify dominant values in order to carry out an overall harmony shift sensitive to tones, tints, and shades.

In this experiment, we manipulate only hue values shifting them to different ranges in the RGB colour wheel according to the nearest detected harmony: monochromatic, analogous, complementary,
We also applied a forced crop that considers the rule of thirds.

Technology for large-scale aesthetic inference is widely available and does not attempt to predict reception of images by single users. NIMA [18] is a neural architecture for image assessment that predicts a distribution of ratings from one to ten. It improves handling of ground-truth ambiguity by optimizing the Earth Mover’s Distance on ordered user score distributions. In addition to the mean user rating, the distribution of NIMA can capture agreement of user ratings. The image alone does not provide all information of the user state that could influence the rating (e.g., memories from the moment the user took the photo, current mood). Figure 2 illustrates that there is a mismatch between the distributions of per-image means and standard deviations when comparing the ground truth to the predictions of NIMA. NIMA predicts the overall reception of an image by users and does not attempt to predict reception of images by single users.

Figure 2: A histogram of the per image mean and standard deviation as calculated on ground truth and as predicted by NIMA, figure from [18].

4 PERCEPTION OF IMAGE ENHANCEMENT

The user study is aimed at gathering qualitative insight into aspects of image enhancements important for user preference. The study compares three approaches: (1) systematically increasing overall colour harmony and improving composition (cf. Section 2), (2) enhancing the images intuitively, carried out by an artist, who restricted the enhancements to the same sort of manipulations that were applied systematically in (1), and (3) the style transfer approach described in [10]. Each approach is used to generate an enhanced image from ten original images from the manual test set of the 2018 Pixel Privacy task, resulting in 30 image pairs. The order of the pairs is randomised.

For each pair the original and enhanced image are randomly assigned to be Image A and Image B. Study participants look at both images and then answer the question “Which image would you prefer to share?” using a 5-point scale running between A and B. Additionally they give qualitative feedback by giving a short elaboration on their preference. The interface allows the user to toggle between image A and image B. Toggling makes the interface more closely resemble the user interfaces in existing applications (e.g., Instagram) and is also intended to eliminate unwanted direct comparisons of the two images. We had access to a group of people with conscious knowledge about images (e.g., photography or computer vision expertise), and, for this preliminary, we selected the study participants (ten in total) from this group. The rationale is that this group would be better able to identify which of their reactions is related to image transformations (as opposed to content) and express their reactions in words.

On average, study participants preferred the original image over the enhanced image. For systematic enhancement (1), enhanced images were preferred in 2 of the 10 cases, compared to 3/10 for intuitive enhancement (2). We identified several high-level categories capturing generalisations in the reasons given by study participants for their image preferences: colours (harmony, cold/warm), composition (ratio, perspective, focus, information, framing), no difference, authenticity (water is not purple). Notable was that for the systematic enhancement, composition change has an effect on the perceived authenticity and image quality (with respect to focus).

5 OUTLOOK

In this paper, we have explored three different views on image enhancement. Aspects from the field of image aesthetics can be systematised for specific image enhancement, making the change more dependent on the content of the image, while not requiring full understanding of the image. For the Pixel Privacy task, machine learning aesthetic assessment does not treat users equally in terms of the error of prediction of their appeal judgements. The Pixel Privacy task could benefit from automatic assessment that treats all users equally in terms of the prediction error of their appeal judgements.

ACKNOWLEDGMENTS

This work is part of the Open Mind research programme, financed by the Netherlands Organisation for Scientific Research (NWO).
REFERENCES


